Inverse problems in imaging science: From classical regularization methods To state of the art Bayesian methods

Ali Mohammad-Djafari
Laboratoire des Signaux et Systèmes,
UMR8506 CNRS-SUPELEC-UNIV PARIS SUD 11
SUPELEC, 91192 Gif-sur-Yvette, France
http://lss.supelec.free.fr

Email: djafari@lss.supelec.fr
http://djafari.free.fr

Content

1. Preliminaries: Direct and indirect observation, errors and probability law, 1D signal, 2D and 3D image, ...
2. Inverse problems examples in imaging science
3. Classical methods: Generalized inversion and Regularization
4. Bayesian approach for inverse problems
5. Prior modeling
   - Gaussian, Generalized Gaussian (GG), Gamma, Beta,
   - Gauss-Markov, GG-Marvov
   - Sparsity enforcing priors (Bernouilli-Gaussian, B-Gamma, Cauchy, Student-t, Laplace)
6. Full Bayesian approach (Estimation of hyperparameters)
7. Hierarchical prior models
8. Bayesian Computation and Algorithms for Hierarchical models
9. Gauss-Markov-Potts family of priors
10. Applications and case studies

Preliminaries: Direct and indirect observation

- Direct observation of a few quantities are possible:
  length, time, electrical charge, number of particles
- For many others, we only can measure them by transforming them (Indirect observation). Example:
  Thermometer transforms variation of temeprature to variation of length.
- Imaging science is a perfect example of indirect observation particularly when we want to see inside of a body from the outside (Computed Tomography)
- When measuring (observing) a quantity, the errors are always present.
- For any quantity (direct or indirect observation) we may define a probability law

Probability law: Discrete and continuous variables

- A quantity can be discrete or continuous
- For discrete value quantities we define a probability distribution

$$P(X=k)=p_k,\; k=1,\cdots,K \quad \text{with} \sum_{k=1}^{K} p_k = 1$$

- For continuous value quantities we define a probability density.

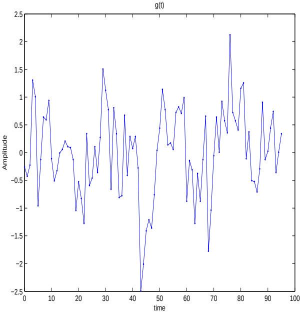
$$P(a < X \leq b) = \int_a^b p(x)\, \mathrm{d}x \quad \text{with} \int_{-\infty}^{+\infty} p(x)\, \mathrm{d}x = 1$$

- For both cases, we may define:
  - Most probable (Mode), Median, Quantiles
  - Regions of high probabilities, ...
  - Expected value (Mean)
  - Variance, Covariance
  - Higher order moments
  - Entropy

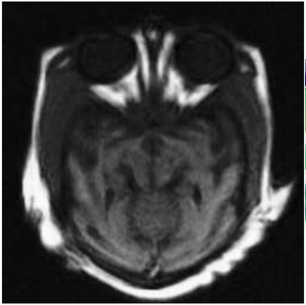
Representation of signals and images

- Signal: $f(t), f(x), f(\nu)$
  - $f(t)$ Variation of temperature in a given position as a function of time t
  - $f(x)$ Variation of temperature as a function of the position x on a line
  - $f(\nu)$ Variation of temperature as a function of the frequency ν
- Image: $f(x,y), f(x,t), f(\nu,t), f(\nu_1,\nu_2)$
  - $f(x,y)$ Distribution of temperature as a function of the position (x,y)
  - $f(x,t)$ Variation of temperature as a function of x and t
  - ...
- 3D, 3D+t, 3D+ν, ... signals: $f(x,y,z), f(x,y,t), f(x,y,z,t)$
  - $f(x,y,z)$ Distribution of temperature as a function of the position (x,y,z)
  - $f(x,y,z,t)$ Variation of temperature as a function of (x,y,z) and t
  - ...

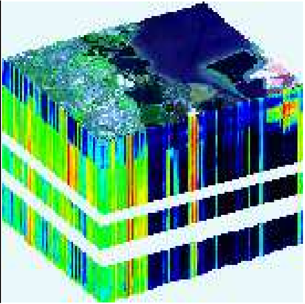
Representation of signals



1D signal



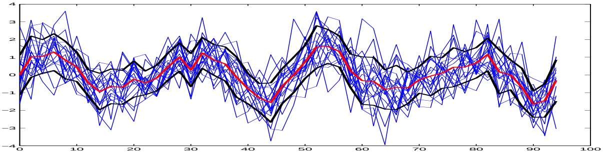
2D signal=image



3D signal

Signals and images

- A signal $f(t)$ can be represented by $p(f(t), t = 0, \cdots, T-1)$



- An image $f(x,y)$ can be represented by $p(f(x,y), (x,y) \in \mathcal{R})$
- Finite domaine observations $\boldsymbol{f} = \{f(t), t = 0, \cdots, T-1\}$
- Image $\boldsymbol{F} = \{f(x,y)\}$ a 2D table or a 1D table $\boldsymbol{f} = \{f(x,y), (x,y) \in \mathcal{R}\}$
- For a vector $\boldsymbol{f}$ we define $p(\boldsymbol{f})$. Then, we can define
  - Most probable value: $\widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f})\}$
  - Expected value : $\boldsymbol{m} = \mathsf{E}\{\boldsymbol{f}\} = \int \boldsymbol{f} p(\boldsymbol{f}) \, \mathsf{d}\boldsymbol{f}$
  - CoVariance matrix: $\boldsymbol{\Sigma} = \mathsf{E}\{(\boldsymbol{f} - \boldsymbol{m})(\boldsymbol{f} - \boldsymbol{m})'\}$
  - Entropy $H = \mathsf{E}\{-\ln p(\boldsymbol{f})\} = -\int p(\boldsymbol{f}) \ln p(\boldsymbol{f}) \, \mathsf{d}\boldsymbol{f}$

2. Inverse problems examples

- Example 1:
  Measuring variation of temperature with a therometer
  - $f(t)$ variation of temperature over time
  - $g(t)$ variation of length of the liquid in thermometer
- Example 2: Seeing outside of a body: Making an image using a camera, a microscope or a telescope
  - $f(x, y)$ real scene
  - $g(x, y)$ observed image
- Example 3: Seeing inside of a body: Computed Tomography usng X rays, US, Microwave, etc.
  - $f(x, y)$ a section of a real 3D body $f(x, y, z)$
  - $g_\phi(r)$ a line of observed radiographe $g_\phi(r, z)$

- Example 1: Deconvolution
- Example 2: Image restoration
- Example 3: Image reconstruction

Measuring variation of temperature with a therometer

- $f(t)$ variation of temperature over time
- $g(t)$ variation of length of the liquid in thermometer

- Forward model: Convolution

$$g(t) = \int f(t')\, h(t - t')\, \mathrm{d}t' + \epsilon(t)$$

$h(t)$: impulse response of the measurement system

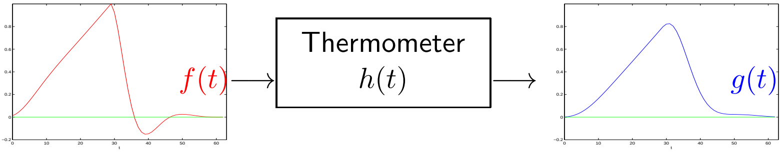
- Inverse problem: Deconvolution

Given the forward model $\mathcal{H}$ (impulse response $h(t)$))
and a set of data $g(t_i), i = 1, \cdots, M$
find $f(t)$

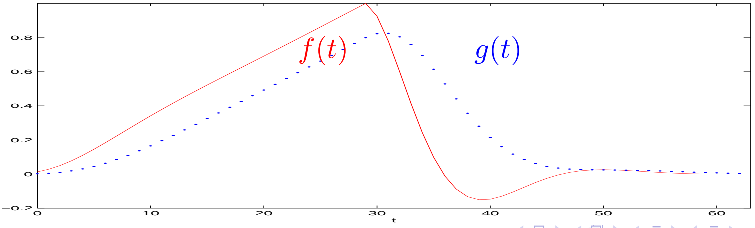
Measuring variation of temperature with a therometer

Forward model: Convolution

$$g(t) = \int f(t')\, h(t-t')\, \mathrm{d}t' + \epsilon(t)$$

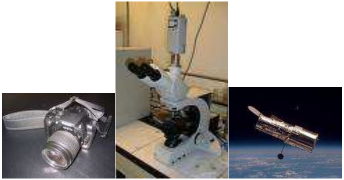


Inversion: Deconvolution



Seeing outside of a body: Making an image with a camera, a microscope or a telescope

- $f(x, y)$ real scene
- $g(x, y)$ observed image



- Forward model: Convolution

$$g(x, y) = \iint f(x', y')\, h(x - x', y - y')\, \mathrm{d}x'\, \mathrm{d}y' + \epsilon(x, y)$$

$h(x, y)$: Point Spread Function (PSF) of the imaging system

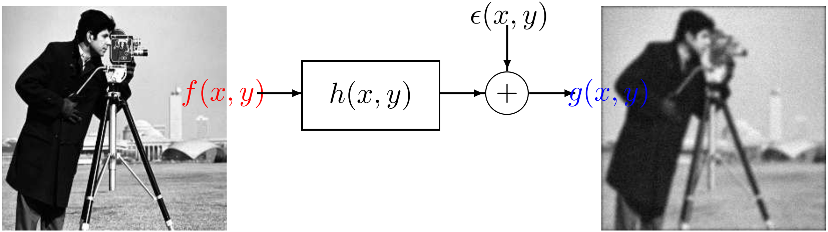
- Inverse problem: Image restoration

Given the forward model $\mathcal{H}$ (PSF $h(x, y)$)
and a set of data $g(x_i, y_i), i = 1, \cdots, M$
find $f(x, y)$

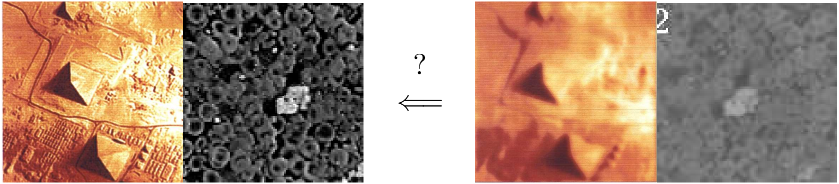
Making an image with an unfocused camera

Forward model: 2D Convolution

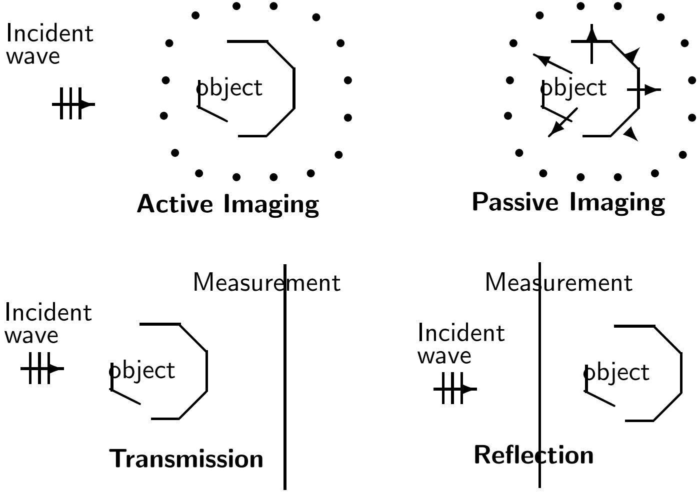
$$g(x,y) = \iint f(x',y')\, h(x-x',y-y')\; \mathrm{d}x'\, \mathrm{d}y' + \epsilon(x,y)$$



Inversion: Image Deconvolution or Restoration



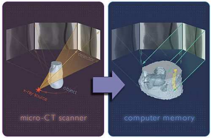
Making an image of the interior of a body
Different imaging systems:



Forward problem: Knowing the object predict the data
Inverse problem: From measured data find the object

Seeing inside of a body: Computed Tomography

- $f(x, y)$ a section of a real 3D body $f(x, y, z)$
- $g_\phi(r)$ a line of observed radiographe $g_\phi(r, z)$



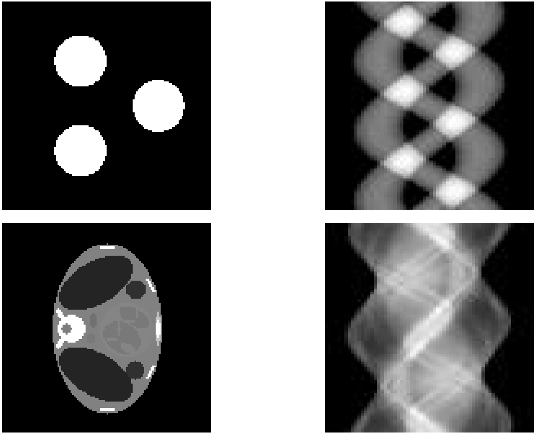
- Forward model:
  Line integrals or Radon Transform

$$\begin{aligned} g_\phi(r) &= \int_{L_{r,\phi}} f(x, y)\, \mathrm{d}l + \epsilon_\phi(r) \\ &= \iint f(x, y)\, \delta(r - x\cos\phi - y\sin\phi)\, \mathrm{d}x\, \mathrm{d}y + \epsilon_\phi(r) \end{aligned}$$

- Inverse problem: Image reconstruction

  Given the forward model $\mathcal{H}$ (Radon Transform) and
  a set of data $g_{\phi_i}(r), i = 1, \cdots, M$
  find $f(x, y)$

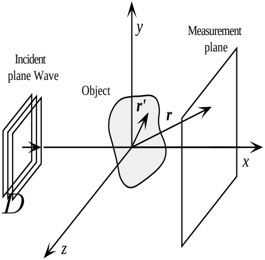
Computed Tomography: Radon Transform



Forward:	$f(x,y)$	$\longrightarrow$	$g(r,\phi)$
Inverse:	$f(x,y)$	$\longleftarrow$	$g(r,\phi)$

Microwave or ultrasound imaging

Measurs: diffracted wave by the object $g(\boldsymbol{r}_i)$
Unknown quantity: $f(\boldsymbol{r}) = k_0^2(n^2(\boldsymbol{r}) - 1)$
Intermediate quantity : $\phi(\boldsymbol{r})$

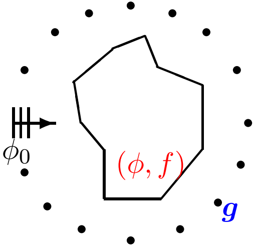


$$g(\boldsymbol{r}_i) = \iint_D G_m(\boldsymbol{r}_i, \boldsymbol{r}')\phi(\boldsymbol{r}')\, f(\boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}', \; \boldsymbol{r}_i \in S$$

$$\phi(\boldsymbol{r}) = \phi_0(\boldsymbol{r}) + \iint_D G_o(\boldsymbol{r}, \boldsymbol{r}')\phi(\boldsymbol{r}')\, f(\boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}', \; \boldsymbol{r} \in D$$

Born approximation $(\phi(\boldsymbol{r}') \simeq \phi_0(\boldsymbol{r}'))$ **):**

$$g(\boldsymbol{r}_i) = \iint_D G_m(\boldsymbol{r}_i, \boldsymbol{r}')\phi_0(\boldsymbol{r}')\, f(\boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}', \; \boldsymbol{r}_i \in S$$



Discretization :

$$\begin{cases} \boldsymbol{g} = \boldsymbol{G}_m \boldsymbol{F} \boldsymbol{\phi} \\ \boldsymbol{\phi} = \boldsymbol{\phi}_0 + \boldsymbol{G}_o \boldsymbol{F} \boldsymbol{\phi} \end{cases} \longrightarrow \begin{cases} \boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) \\ \text{with } \; \boldsymbol{F} = \mathrm{diag}(\boldsymbol{f}) \\ \boldsymbol{H}(\boldsymbol{f}) = \boldsymbol{G}_m \boldsymbol{F} (\boldsymbol{I} - \boldsymbol{G}_o \boldsymbol{F})^{-1} \boldsymbol{\phi}_0 \end{cases}$$

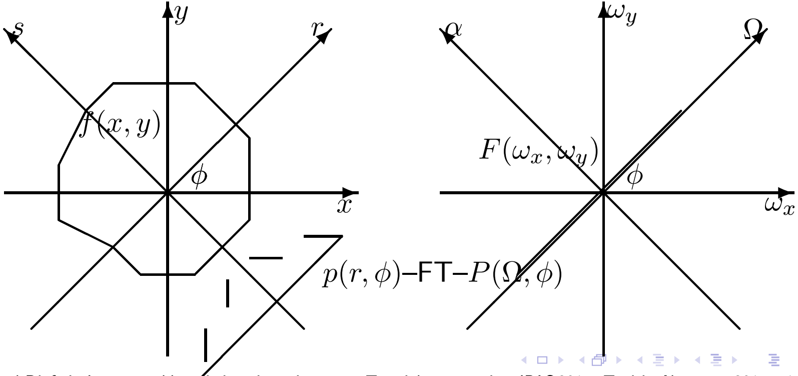
Fourier Synthesis in X ray Tomography

$$g(r,\phi) = \iint f(x,y)\, \delta(r - x\cos\phi - y\sin\phi)\, \mathrm{d}x\, \mathrm{d}y$$

$$G(\Omega,\phi) = \int g(r,\phi) \exp\left[-j\Omega r\right]\, \mathrm{d}r$$

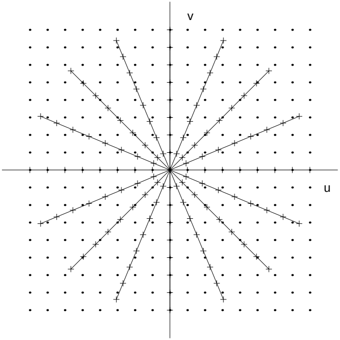
$$F(\omega_x,\omega_y) = \iint f(x,y) \exp\left[-j\omega_x x, \omega_y y\right]\, \mathrm{d}x\, \mathrm{d}y$$

$$F(\omega_x,\omega_y) = G(\Omega,\phi) \quad \text{for} \quad \omega_x = \Omega\cos\phi \quad \text{and} \quad \omega_y = \Omega\sin\phi$$

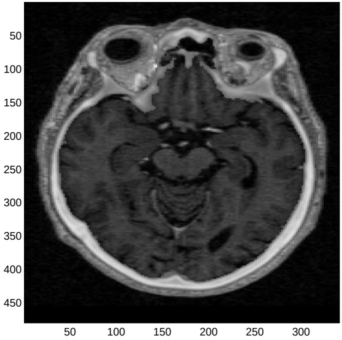


Fourier Synthesis in X ray tomography

$$G(\omega_x, \omega_y) = \iint f(x,y) \exp\left[-j\left(\omega_x x + \omega_y y\right)\right] \, \mathrm{d}x \, \mathrm{d}y$$

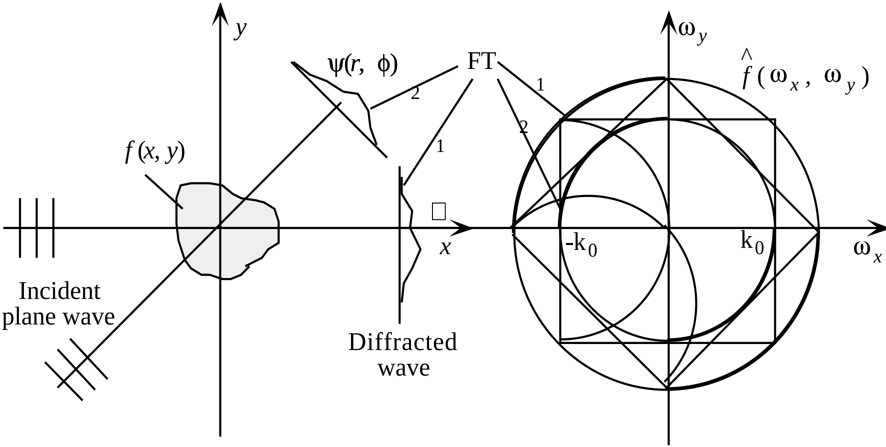


? $\Longrightarrow$



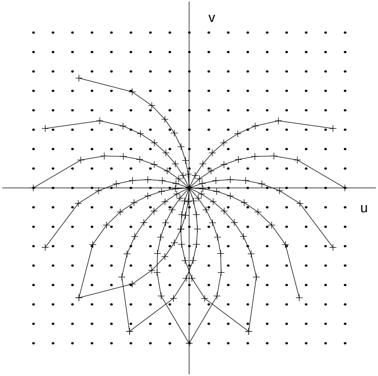
Forward problem: Given $f(x,y)$ compute $G(\omega_x, \omega_y)$
Inverse problem: Given $G(\omega_x, \omega_y)$ on those lines estimate $f(x,y)$

Fourier Synthesis in Diffraction tomography

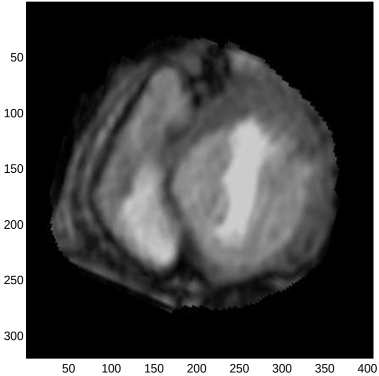


Fourier Synthesis in Diffraction tomography

$$G(\omega_x, \omega_y) = \iint f(x, y) \exp\left[-j\left(\omega_x x + \omega_y y\right)\right] \, \mathrm{d}x \, \mathrm{d}y$$



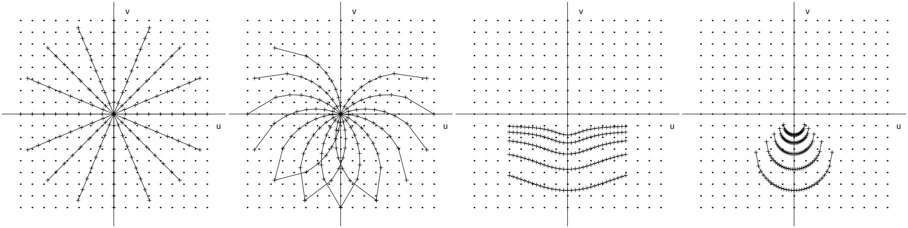
$$? \\ \Longrightarrow$$



Forward problem: Given $f(x, y)$ compute $G(\omega_x, \omega_y)$
Inverse problem : Given $G(\omega_x, \omega_y)$ on those semi cercles estimate $f(x, y)$

Fourier Synthesis in different imaging systems

$$G(\omega_x, \omega_y) = \iint f(x, y) \exp\left[-j\left(\omega_x x + \omega_y y\right)\right] \, dx \, dy$$



X ray Tomography | Diffraction | Eddy current | SAR & Radar

Forward problem: Given $f(x, y)$ compute $G(\omega_x, \omega_y)$
Inverse problem : Given $G(\omega_x, \omega_y)$ on those algebraic lines, cercles or curves, estimate $f(x, y)$

Invers Problems: other examples and applications

- X ray, Gamma ray Computed Tomography (CT)
- Microwave and ultrasound tomography
- Positron emission tomography (PET)
- Magnetic resonance imaging (MRI)
- Photoacoustic imaging
- Radio astronomy
- Geophysical imaging
- Non Destructive Evaluation (NDE) and Testing (NDT) techniques in industry
- Hyperspectral imaging
- Earth observation methods (Radar, SAR, IR, ...)
- Survey and tracking in security systems

3. General formulation of inverse problems and classical methods

- General non linear inverse problems:

$$g(\boldsymbol{s}) = [\mathcal{H} f(\boldsymbol{r})](\boldsymbol{s}) + \epsilon(\boldsymbol{s}), \quad \boldsymbol{r} \in \mathcal{R}, \quad \boldsymbol{s} \in \mathcal{S}$$

- Linear models:

$$g(\boldsymbol{s}) = \int f(\boldsymbol{r})\, h(\boldsymbol{r}, \boldsymbol{s})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{s})$$

  If $h(\boldsymbol{r}, \boldsymbol{s}) = h(\boldsymbol{r} - \boldsymbol{s}) \longrightarrow$ Convolution.

- Discrete data:

$$g(\boldsymbol{s}_i) = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, f(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{s}_i), \quad i = 1, \cdots, m$$

- Inversion: Given the forward model $\mathcal{H}$ and the data $\boldsymbol{g} = \{g(\boldsymbol{s}_i), i = 1, \cdots, m)\}$ estimate $f(\boldsymbol{r})$
- Well-posed and Ill-posed problems (Hadamard): existance, uniqueness and stability
- Need for prior information

Inverse problems: Discretization

$$g(\boldsymbol{s}_i) = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, f(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{s}_i), \quad i = 1, \cdots, M$$

- $f(\boldsymbol{r})$ is assumed to be well approximated by

$$f(\boldsymbol{r}) \simeq \sum_{j=1}^{N} f_j\, b_j(\boldsymbol{r})$$

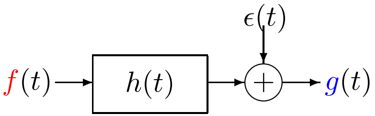
with $\{b_j(\boldsymbol{r})\}$ a basis or any other set of known functions

$$g(\boldsymbol{s}_i) = g_i \simeq \sum_{j=1}^{N} f_j \int h(\boldsymbol{s}_i, \boldsymbol{r})\, b_j(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}, \quad i = 1, \cdots, M$$

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \;\text{ with }\; H_{ij} = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, b_j(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}$$

- $\boldsymbol{H}$ is huge dimensional
- LS solution : $\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{Q(\boldsymbol{f})\}$ with
  $Q(\boldsymbol{f}) = \sum_i |g_i - [\boldsymbol{H}\boldsymbol{f}]_i|^2 = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2$
  does not give satisfactory result.

Convolution: Discretization



$$g(t) = \int f(t')\, h(t-t')\, \mathrm{d}t' + \epsilon(t) = \int h(t')\, f(t-t')\, \mathrm{d}t' + \epsilon(t)$$

- The signals $f(t)$, $g(t)$, $h(t)$ are discretized with the same sampling period $\Delta T = 1$,
- The impulse response is finite (FIR) : $h(t) = 0$, for t such that $t < -q\Delta T$ or $\forall t > p\Delta T$.

$$g(m) = \sum_{k=-q}^{p} h(k)\, f(m-k) + \epsilon(m), \quad m = 0, \cdots, M$$

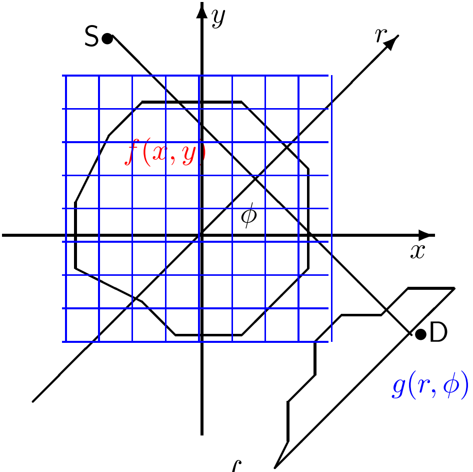
Convolution: Discretized matrix vector form

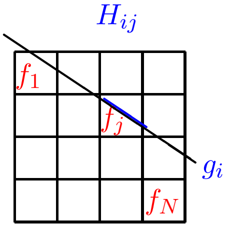
- If system is causal ($q = 0$) we obtain

$$\begin{bmatrix} g(0) \\ g(1) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ g(M) \end{bmatrix} = \begin{bmatrix} h(p) & \cdots & h(0) & 0 & \cdots & \cdots & 0 \\ 0 & & & & & & \vdots \\ \vdots & & & & & & \vdots \\ \vdots & & h(p) & \cdots & h(0) & & \vdots \\ \vdots & & & & & & \vdots \\ \vdots & & & & & & 0 \\ 0 & \cdots & \cdots & 0 & h(p) & \cdots & h(0) \end{bmatrix} \begin{bmatrix} f(-p) \\ \vdots \\ f(0) \\ f(1) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ f(M) \end{bmatrix}$$

- $\boldsymbol{g}$ is a $(M+1)$-dimensional vector,
- $\boldsymbol{f}$ has dimension $M+p+1$,
- $\boldsymbol{h} = [h(p), \cdots, h(0)]$ has dimension $(p+1)$
- $\boldsymbol{H}$ has dimensions $(M+1) \times (M+p+1)$.

Discretization of Radon Transfrom in CT





$$f(x,y) = \sum_j f_j \, b_j(x,y)$$

$$b_j(x,y) = \begin{cases} 1 & \text{if } (x,y) \in \text{ pixel } j \\ 0 & \text{else} \end{cases}$$

$$g(r,\phi) = \int_L f(x,y) \, \mathrm{d}l$$

$$g_i = \sum_{j=1}^{N} H_{ij} \, f_j + \epsilon_i$$

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

Inverse problems: Deterministic methods

Data matching

- Observation model
  $g_i = h_i(\boldsymbol{f}) + \epsilon_i, \quad i = 1, \ldots, M \longrightarrow \boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}$
- Misatch between data and output of the model $\Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f}))$

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{\Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f}))\}$$

- Examples:

– LS $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \|\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})\|^2 = \sum_i |g_i - h_i(\boldsymbol{f})|^2$

– L_p $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \|\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})\|^p = \sum_i |g_i - h_i(\boldsymbol{f})|^p, \quad 1 < p < 2$

– KL $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \sum_i g_i \ln \frac{g_i}{h_i(\boldsymbol{f})}$

- In general, does not give satisfactory results for inverse problems.

Inverse problems: Regularization theory

Inverse problems = Ill posed problems

$\longrightarrow$ Need for prior information

Functional space (Tikhonov):

$$g = \mathcal{H}(f) + \epsilon \longrightarrow J(f) = ||g - \mathcal{H}(f)||_2^2 + \lambda ||\mathcal{D} f||_2^2$$

Finite dimensional space (Philips & Towmey): $\boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}$

- Minimum norme LS (MNLS): $J(\boldsymbol{f}) = ||\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})||^2 + \lambda ||\boldsymbol{f}||^2$
- Classical regularization: $J(\boldsymbol{f}) = ||\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})||^2 + \lambda ||\boldsymbol{D}\boldsymbol{f}||^2$
- More general regularization:

$$J(\boldsymbol{f}) = \mathcal{Q}(\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})) + \lambda \Omega(\boldsymbol{D}\boldsymbol{f})$$

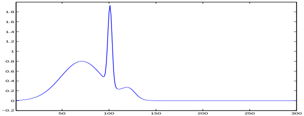
or

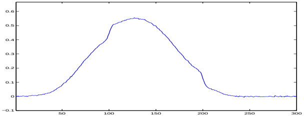
$$J(\boldsymbol{f}) = \Delta_1(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) + \lambda \Delta_2(\boldsymbol{f}, \boldsymbol{f}_\infty)$$

Limitations:

- Errors are considered implicitly white and Gaussian
- Limited prior information on the solution
- Lack of tools for the determination of the hyperparameters

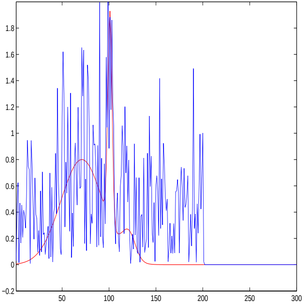
Deconvolution example



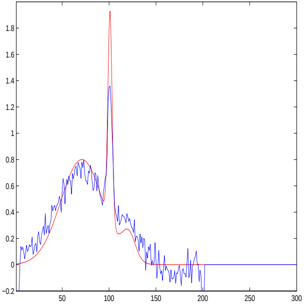


Forward: $f(t) \longrightarrow g(t) = h(t) * f(t) + \epsilon(t)$

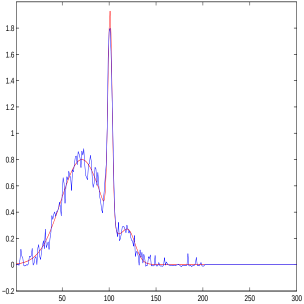
Inverse: $\widehat{f}(t) \longleftarrow g(t)$



Inverse Filtering



Wiener



Regularization

Inversion: Probabilistic methods

Taking account of errors and uncertainties $\longrightarrow$ Probability theory

- Maximum Likelihood (ML)
- Minimum Inaccuracy (MI)
- Probability Distribution Matching (PDM)
- Maximum Entropy (ME) and Information Theory (IT)
- Bayesian Inference (BAYES)

Advantages:

- Explicit account of the errors and noise
- A large class of priors via explicit or implicit modeling
- A coherent approach to combine information content of the data and priors

Limitations:

- Practical implementation and cost of calculation

4. Bayesian inference for inverse problems

$$\mathcal{M}: \qquad \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

- Observation model $\mathcal{M}$ + Hypothesis on the noise $\boldsymbol{\epsilon} \longrightarrow$
  $p(\boldsymbol{g}|\boldsymbol{f};\mathcal{M}) = p_{\boldsymbol{\epsilon}}(\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f})$
- A priori information $\quad p(\boldsymbol{f}|\mathcal{M})$
- Bayes : $\quad p(\boldsymbol{f}|\boldsymbol{g};\mathcal{M}) = \dfrac{p(\boldsymbol{g}|\boldsymbol{f};\mathcal{M})\, p(\boldsymbol{f}|\mathcal{M})}{p(\boldsymbol{g}|\mathcal{M})}$

Link with regularization :

Maximum A Posteriori (MAP) :

$$\begin{aligned}
\widehat{\boldsymbol{f}} &= \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g})\} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{g}|\boldsymbol{f})\; p(\boldsymbol{f})\} \\
&= \arg\min_{\boldsymbol{f}} \{-\ln p(\boldsymbol{g}|\boldsymbol{f}) - \ln p(\boldsymbol{f})\}
\end{aligned}$$

with $\quad Q(\boldsymbol{g}, \boldsymbol{H}\boldsymbol{f}) = -\ln p(\boldsymbol{g}|\boldsymbol{f}) \quad$ and $\quad \lambda\Omega(\boldsymbol{f}) = -\ln p(\boldsymbol{f})$

Case of linear models and Gaussian priors

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

- Hypothesis on the noise: $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_\epsilon^2 \boldsymbol{I})$

$$p(\boldsymbol{g}|\boldsymbol{f}) \propto \exp\left[-\frac{1}{2\sigma_\epsilon^2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2\right]$$

- Hypothesis on $\boldsymbol{f}$: $\boldsymbol{f} \sim \mathcal{N}(0, \sigma_f^2 \boldsymbol{I})$

$$p(\boldsymbol{f}) \propto \exp\left[-\frac{1}{2\sigma_f^2}\|\boldsymbol{f}\|^2\right]$$

- A posteriori:

$$p(\boldsymbol{f}|\boldsymbol{g}) \propto \exp\left[-\frac{1}{2\sigma_\epsilon^2}\left(\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \frac{\sigma_\epsilon^2}{\sigma_f^2}\|\boldsymbol{f}\|^2\right)\right]$$

- MAP : $\widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g})\} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\}$

  with $J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\|\boldsymbol{f}\|^2, \qquad \lambda = \frac{\sigma_\epsilon^2}{\sigma_f^2}$

- Advantage : characterization of the solution

$\boldsymbol{f}|\boldsymbol{g} \sim \mathcal{N}(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{P}})$ with $\widehat{\boldsymbol{f}} = \left(\boldsymbol{H}^t\boldsymbol{H} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{H}^t\boldsymbol{g}$, $\widehat{\boldsymbol{P}} = \sigma_\epsilon^2\left(\boldsymbol{H}^t\boldsymbol{H} + \lambda\boldsymbol{I}\right)^{-1}$

MAP estimation with other priors:

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\} \quad \text{with} \quad J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\Omega(\boldsymbol{f})$$

Separable priors:

- Gaussian:
  $p(f_j) \propto \exp\left[-\alpha|f_j|^2\right] \longrightarrow \Omega(\boldsymbol{f}) = \|\boldsymbol{f}\|^2 = \alpha \sum_j |f_j|^2$
- Gamma: $p(f_j) \propto f_j^{\alpha} \exp\left[-\beta f_j\right] \longrightarrow \Omega(\boldsymbol{f}) = \alpha \sum_j \ln f_j + \beta f_j$
- Beta:
  $p(f_j) \propto f_j^{\alpha}(1 - f_j)^{\beta} \longrightarrow \Omega(\boldsymbol{f}) = \alpha \sum_j \ln f_j + \beta \sum_j \ln(1 - f_j)$
- Generalized Gaussian: $p(f_j) \propto \exp\left[-\alpha|f_j|^p\right], \quad 1 < p < 2 \longrightarrow \quad \Omega(\boldsymbol{f}) = \alpha \sum_j |f_j|^p,$

Markovian models:

$$p(f_j|\boldsymbol{f}) \propto \exp\left[-\alpha \sum_{i \in N_j} \phi(f_j, f_i)\right] \longrightarrow \quad \Omega(\boldsymbol{f}) = \alpha \sum_j \sum_{i \in N_j} \phi(f_j, f_i),$$

MAP estimation and Compressed Sensing

$$\begin{cases} \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \\ \boldsymbol{f} = \boldsymbol{W}\boldsymbol{z} \end{cases}$$

- $\boldsymbol{W}$ a code book matrix, $\boldsymbol{z}$ coefficients
- Gaussian:

$$p(\boldsymbol{z}) = \mathcal{N}(0, \sigma_z^2 \boldsymbol{I}) \propto \exp\left[-\frac{1}{2\sigma_z^2} \sum_j |z_j|^2\right]$$
$$J(\boldsymbol{z}) = -\ln p(\boldsymbol{z}|\boldsymbol{g}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{W}\boldsymbol{z}\|^2 + \lambda \sum_j |z_j|^2$$

- Generalized Gaussian (sparsity, $\beta = 1$):

$$p(\boldsymbol{z}) \propto \exp\left[-\lambda \sum_j |z_j|^\beta\right]$$
$$J(\boldsymbol{z}) = -\ln p(\boldsymbol{z}|\boldsymbol{g}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{W}\boldsymbol{z}\|^2 + \lambda \sum_j |z_j|^\beta$$

- $\boldsymbol{z} = \arg\min_{\boldsymbol{z}} \{J(\boldsymbol{z})\} \longrightarrow \widehat{\boldsymbol{f}} = \boldsymbol{W}\widehat{\boldsymbol{z}}$

Bayesian Estimation: Two simple priors

- Example 1: Linear Gaussian case:

$$\left\{\begin{array}{l} p(\boldsymbol{g}|\boldsymbol{f},\theta_1)=\mathcal{N}(\boldsymbol{H}\boldsymbol{f},\theta_1\boldsymbol{I}) \\ p(\boldsymbol{f}|\theta_2)=\mathcal{N}(0,\theta_2\boldsymbol{I}) \end{array}\right. \longrightarrow p(\boldsymbol{f}|\boldsymbol{g},\boldsymbol{\theta})=\mathcal{N}(\widehat{\boldsymbol{f}},\widehat{\boldsymbol{P}})$$

with

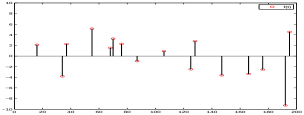
$$\left\{\begin{array}{l} \widehat{\boldsymbol{f}}=(\boldsymbol{H}'\boldsymbol{H}+\lambda\boldsymbol{I})^{-1}\boldsymbol{H}'\boldsymbol{g} \\ \widehat{\boldsymbol{P}}=\theta_1(\boldsymbol{H}'\boldsymbol{H}+\lambda\boldsymbol{I})^{-1}, \quad \lambda=\frac{\theta_1}{\theta_2} \end{array}\right.$$

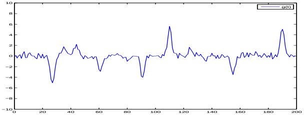
$$\widehat{\boldsymbol{f}}=\arg\min_{\boldsymbol{f}}\{J(\boldsymbol{f})\} \text{ with } J(\boldsymbol{f})=\|\boldsymbol{g}-\boldsymbol{H}\boldsymbol{f}\|_2^2+\lambda\|\boldsymbol{f}\|_2^2$$

- Example 2: Double Exponential prior & MAP:

$$\widehat{\boldsymbol{f}}=\arg\min_{\boldsymbol{f}}\{J(\boldsymbol{f})\} \text{ with } J(\boldsymbol{f})=\|\boldsymbol{g}-\boldsymbol{H}\boldsymbol{f}\|_2^2+\lambda\|\boldsymbol{f}\|_1$$

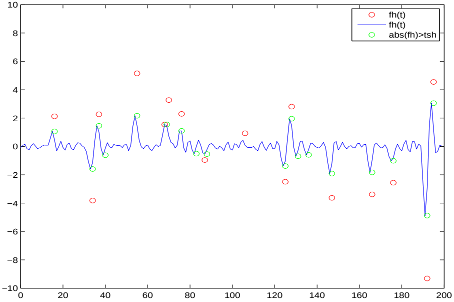
Deconvolution example



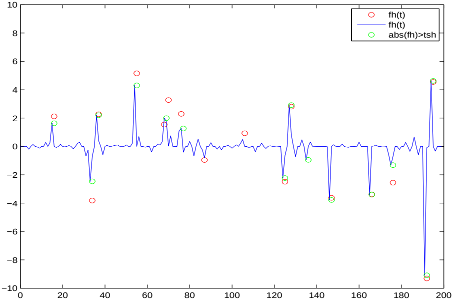


Forward: $f(t) \longrightarrow g(t) = h(t) * f(t) + \epsilon(t)$

Inverse: $\widehat{f}(t) \longleftarrow g(t)$



Quadratic Reg. (Gaussian)



L_1 Reg. (Laplace)

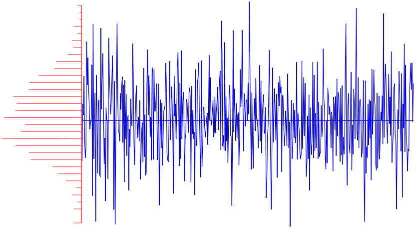
Main advantages of the Bayesian approach

- MAP = Regularization
- Posterior mean ? Marginal MAP ?
- More information in the posterior law than only its mode or its mean
- Meaning and tools for estimating hyper parameters
- Meaning and tools for model selection
- More specific and specialized priors, particularly through the hidden variables
- More computational tools:
  - Expectation-Maximization for computing the maximum likelihood parameters
  - MCMC for posterior exploration
  - Variational Bayes for analytical computation of the posterior marginals
  - ...

Two main steps in the Bayesian approach

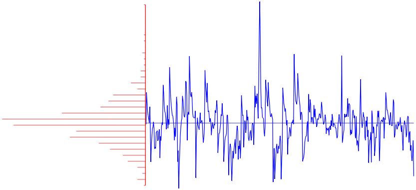
- Prior modeling
  - Separable:
    Gaussian, Gamma,
    Sparsity enforcing: Generalized Gaussian, mixture of Gaussians, mixture of Gammas, ...
  - Markovian:
    Gauss-Markov, GGM, ...
  - Markovian with hidden variables
    (contours, region labels)
- Choice of the estimator and computational aspects
  - MAP, Posterior mean, Marginal MAP
  - MAP needs optimization algorithms
  - Posterior mean needs integration methods
  - Marginal MAP and Hyperparameter estimation need integration and optimization
  - Approximations:
    - Gaussian approximation (Laplace)
    - Numerical exploration MCMC
    - Variational Bayes (Separable approximation)

5. Prior modeling of signals



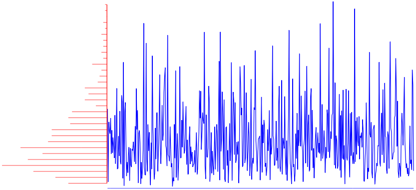
Gaussian

$p(f_j) \propto \exp\left[-\alpha |f_j|^2\right]$



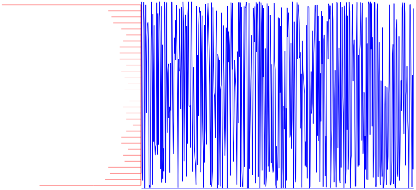
Generalized Gaussian

$p(f_j) \propto \exp\left[-\alpha |f_j|^p\right], \quad 1 \leq p \leq 2$



Gamma

$p(f_j) \propto f_j^\alpha \exp\left[-\beta f_j\right]$

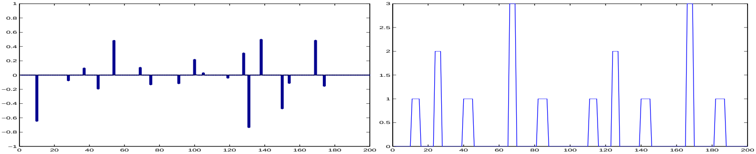


Beta

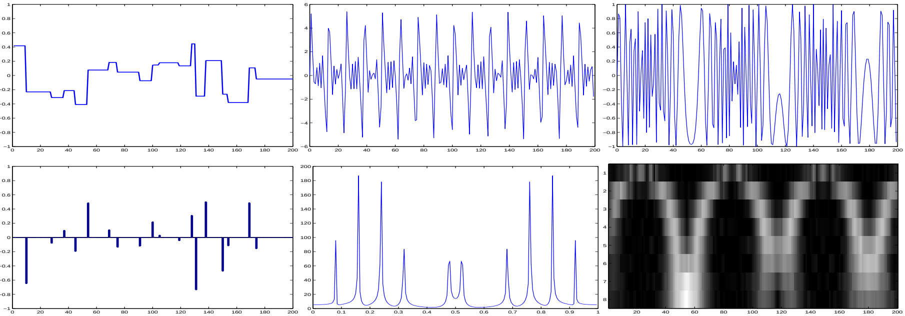
$p(f_j) \propto f_j^\alpha (1 - f_j)^\beta$

Sparsity enforcing prior models

- Sparse signals: Direct sparsity



- Sparse signals: Sparsity in a Transform domaine



Sparsity enforcing prior models

- Simple heavy tailed models:
  - Generalized Gaussian, Double Exponential
  - Student-t, Cauchy
  - Elastic net

  - Symmetric Weibull, Symmetric Rayleigh
  - Generalized hyperbolic

- Hierarchical mixture models:
  - Mixture of Gaussians
  - Bernoulli-Gaussian

  - Mixture of Gammas
  - Bernoulli-Gamma
  - Mixture of Dirichlet
  - Bernoulli-Multinomial

Simple heavy tailed models

- Generalized Gaussian, Double Exponential

$$p(\boldsymbol{f}|\gamma,\beta) = \prod_j \mathcal{GG}(f_j|\gamma,\beta) \propto \exp\left[-\gamma \sum_j |f_j|^\beta\right]$$

$\beta = 1$ Double exponential or Laplace.
$0 < \beta \leq 1$ are of great interest for sparsity enforcing.

- Student-t and Cauchy models

$$p(\boldsymbol{f}|\nu) = \prod_j \mathcal{S}t(f_j|\nu) \propto \exp\left[-\frac{\nu+1}{2} \sum_j \log\left(1 + f_j^2/\nu\right)\right]$$

Cauchy model is obtained when $\nu = 1$.

Mixture models

- Mixture of two Gaussians (MoG2) model

$$p(\boldsymbol{f}|\lambda, v_1, v_0) = \prod_j \left(\lambda \mathcal{N}(f_j|0, v_1) + (1-\lambda)\mathcal{N}(f_j|0, v_0)\right)$$

- Bernoulli-Gaussian (BG) model

$$p(\boldsymbol{f}|\lambda, v) = \prod_j p(f_j) = \prod_j \left(\lambda \mathcal{N}(f_j|0, v) + (1-\lambda)\delta(f_j)\right)$$

- Mixture of Gammas

$$p(\boldsymbol{f}|\lambda, v_1, v_0) = \prod_j \left(\lambda \mathcal{G}(f_j|\alpha_1, \beta_1) + (1-\lambda)\mathcal{G}(f_j|\alpha_2, \beta_2)\right)$$

- Bernoulli-Gamma model

$$p(\boldsymbol{f}|\lambda, \alpha, \beta) = \prod_j \left[\lambda \mathcal{G}(f_j|\alpha, \beta) + (1-\lambda)\delta(f_j)\right]$$

6. Full Bayesian approach

$$\mathcal{M}: \qquad \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

- Forward & errors model: $\longrightarrow p(\boldsymbol{g}|\boldsymbol{f}, \boldsymbol{\theta}_1; \mathcal{M})$
- Prior models $\longrightarrow p(\boldsymbol{f}|\boldsymbol{\theta}_2; \mathcal{M})$
- Hyperparameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \longrightarrow p(\boldsymbol{\theta}|\mathcal{M})$
- Bayes: $\longrightarrow p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M}) = \frac{p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{\theta};\mathcal{M})\; p(\boldsymbol{f}|\boldsymbol{\theta};\mathcal{M})\; p(\boldsymbol{\theta}|\mathcal{M})}{p(\boldsymbol{g}|\mathcal{M})}$
- Joint MAP: $\qquad (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\theta}}) = \arg\max_{(\boldsymbol{f},\boldsymbol{\theta})} \{p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\}$
- Marginalization 1:

$$p(\boldsymbol{f}|\boldsymbol{g}; \mathcal{M}) = \iint p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{\theta} \longrightarrow \widehat{\boldsymbol{f}}$$

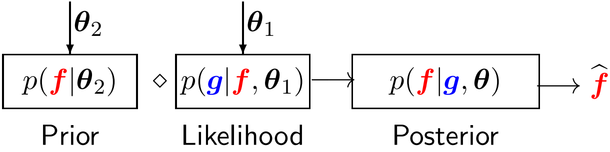
- Marginalization 2:

$$p(\boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M}) = \iint p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{f} \longrightarrow \widehat{\boldsymbol{\theta}}$$

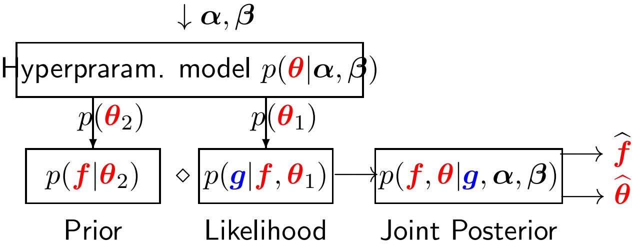
- Approximate $p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})$ by a separable one $q(\boldsymbol{f}, \boldsymbol{\theta}) = q_1(\boldsymbol{f}) q_2(\boldsymbol{\theta})$ and then use them separately to find $\widehat{\boldsymbol{f}}$ and $\widehat{\boldsymbol{\theta}}$.

Summary of Bayesian estimation 1

- Simple Bayesian Model and Estimation



- Full Bayesian Model and Hyperparameter Estimation



Summary of Bayesian estimation 2

- Marginalization 1

$$\boxed{p(\boldsymbol{f},\boldsymbol{\theta}|\boldsymbol{g})} \longrightarrow \boxed{p(\boldsymbol{\theta}|\boldsymbol{g})} \longrightarrow \widehat{\boldsymbol{f}}$$

Joint Posterior   Marginalize over $\boldsymbol{\theta}$

- Marginalization 2

$$\boxed{p(\boldsymbol{f},\boldsymbol{\theta}|\boldsymbol{g})} \longrightarrow \boxed{p(\boldsymbol{\theta}|\boldsymbol{g})} \longrightarrow \widehat{\boldsymbol{\theta}} \longrightarrow \boxed{p(\boldsymbol{f}|\widehat{\boldsymbol{\theta}},\boldsymbol{g})} \longrightarrow \widehat{\boldsymbol{f}}$$

Joint Posterior   Marginalize over $\boldsymbol{f}$

- Variational Bayesian Approximation

$$\boxed{p(\boldsymbol{f},\boldsymbol{\theta}|\boldsymbol{g})} \longrightarrow \boxed{\text{Variational Bayesian Approximation}} \begin{array}{l} \longrightarrow q_1(\boldsymbol{f}) \longrightarrow \widehat{\boldsymbol{f}} \\ \longrightarrow q_2(\boldsymbol{\theta}) \longrightarrow \widehat{\boldsymbol{\theta}} \end{array}$$

Variational Bayesian Approximation

- Full Bayesian: $p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f}, \boldsymbol{\theta}_1)\, p(\boldsymbol{f}|\boldsymbol{\theta}_2)\, p(\boldsymbol{\theta})$
- Approximate $p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g})$ by $q(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}) = q_1(\boldsymbol{f}|\boldsymbol{g})\, q_2(\boldsymbol{\theta}|\boldsymbol{g})$ and then continue computations.
- Criterion $\text{KL}(q(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}) : p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}))$
- $\text{KL}(q : p) = \int\int q \ln q/p = \int\int q_1 q_2 \ln \frac{q_1 q_2}{p}$
- Iterative algorithm $q_1 \longrightarrow q_2 \longrightarrow q_1 \longrightarrow q_2, \cdots$

$$\left\{ \begin{array}{ll} \widehat{q}_1(\boldsymbol{f}) & \propto \exp\left[\langle \ln p(\boldsymbol{g}, \boldsymbol{f}, \boldsymbol{\theta}; \mathcal{M})\rangle_{\widehat{q}_2(\boldsymbol{\theta})}\right] \\ \widehat{q}_2(\boldsymbol{\theta}) & \propto \exp\left[\langle \ln p(\boldsymbol{g}, \boldsymbol{f}, \boldsymbol{\theta}; \mathcal{M})\rangle_{\widehat{q}_1(\boldsymbol{f})}\right] \end{array} \right.$$

$$\boxed{p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g})} \longrightarrow \boxed{\text{Variational Bayesian Approximation}} \begin{array}{l} \longrightarrow q_1(\boldsymbol{f}) \longrightarrow \widehat{\boldsymbol{f}} \\ \longrightarrow q_2(\boldsymbol{\theta}) \longrightarrow \widehat{\boldsymbol{\theta}} \end{array}$$

7. Hierarchical models and hidden variables

- All the mixture models and some of simple models can be modeled via hidden variables $\boldsymbol{z}$.

$$p(f) = \sum_{k=1}^{K} \alpha_k p_k(f) \longrightarrow \begin{cases} p(f|z=k) = p_k(f), \\ P(z=k) = \alpha_k, \quad \sum_k \alpha_k = 1 \end{cases}$$

- Example 2: Student-t model

$$\mathcal{S}t(f|\nu) \propto \exp\left[-\frac{\nu+1}{2}\log\left(1+f^2/\nu\right)\right]$$

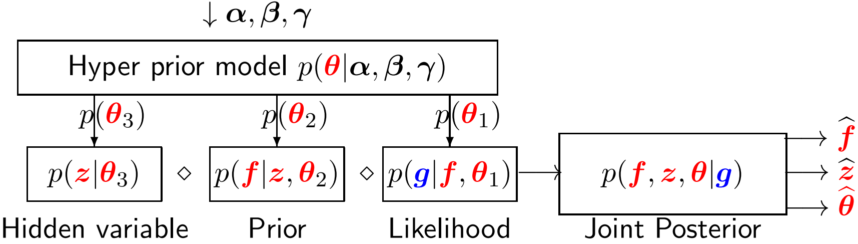
- Infinite mixture

$$\mathcal{S}t(f|\nu) \propto= \int_0^\infty \mathcal{N}(f|,0,1/z)\,\mathcal{G}(z|\alpha,\beta)\,\mathrm{d}z, \quad \text{with } \alpha=\beta=\nu/2$$

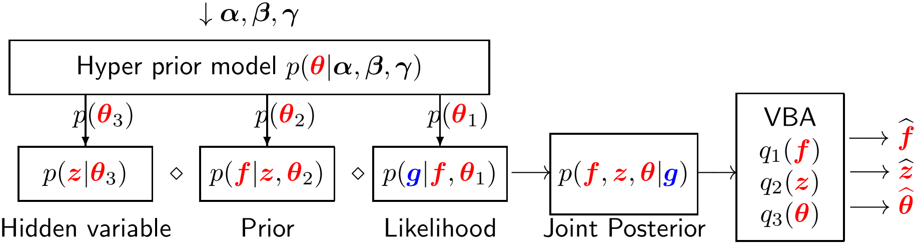
$$\begin{cases} p(\boldsymbol{f}|\boldsymbol{z}) & = \prod_j p(f_j|z_j) = \prod_j \mathcal{N}(f_j|0,1/z_j) \propto \exp\left[-\frac{1}{2}\sum_j z_j f_j^2\right] \\ p(\boldsymbol{z}|\alpha,\beta) & = \prod_j \mathcal{G}(z_j|\alpha,\beta) \propto \prod_j {z_j}^{(\alpha-1)} \exp\left[-\beta z_j\right] \\ & \propto \exp\left[\sum_j (\alpha-1)\ln z_j - \beta z_j\right] \end{cases}$$

Summary of Bayesian estimation 3

- Full Bayesian Hierarchical Model with Hyperparameter Estimation



- Full Bayesian Hierarchical Model and Variational Approximation



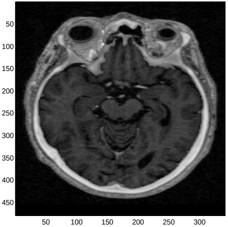
8. Bayesian Computation and Algorithms for Hierarchical models

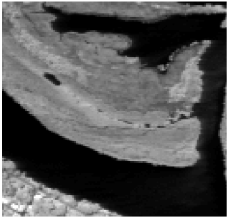
- Often, the expression of $p(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta}|\boldsymbol{g})$ is complex.
- Its optimization (for Joint MAP) or
  its marginalization or integration (for Marginal MAP or PM)
  is not easy
- Two main techniques:
  MCMC and Variational Bayesian Approximation (VBA)
- MCMC:
  Needs the expressions of the conditionals
  $p(\boldsymbol{f}|\boldsymbol{z}, \boldsymbol{\theta}, \boldsymbol{g})$, $p(\boldsymbol{z}|\boldsymbol{f}, \boldsymbol{\theta}, \boldsymbol{g})$, and $p(\boldsymbol{\theta}|\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{g})$
- VBA: Approximate $p(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta}|\boldsymbol{g})$ by a separable one

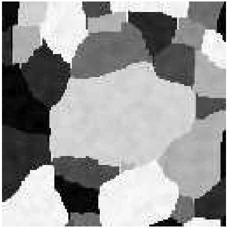
$$q(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta}|\boldsymbol{g}) = q_1(\boldsymbol{f})\, q_2(\boldsymbol{z})\, q_3(\boldsymbol{\theta})$$

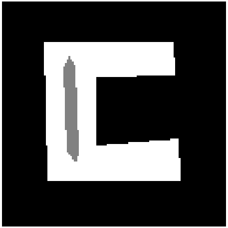
  and do any computations with these separable ones.

Which images I am looking for?

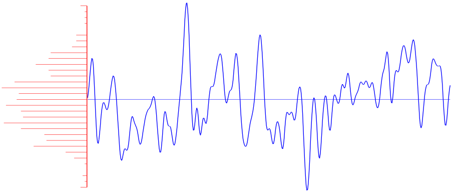




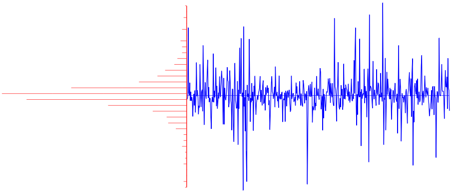




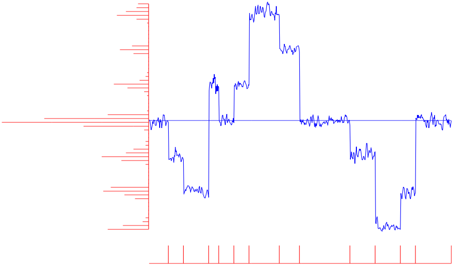
Which image I am looking for?



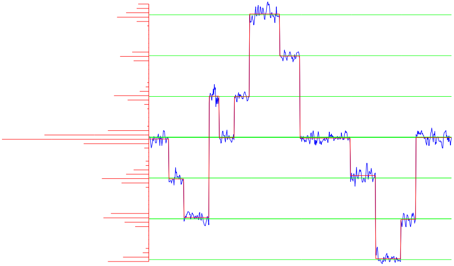
Gauss-Markov



Generalized GM

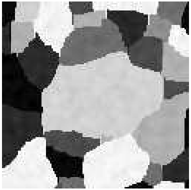


Piecewize Gaussian

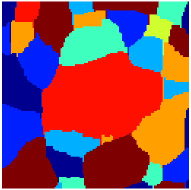


Mixture of GM

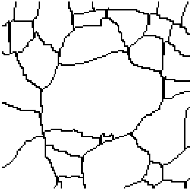
9. Gauss-Markov-Potts prior models for images



$f(\boldsymbol{r})$



$z(\boldsymbol{r})$



$c(\boldsymbol{r}) = 1 - \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))$

$$p(f(\boldsymbol{r})|z(\boldsymbol{r}) = k, m_k, v_k) = \mathcal{N}(m_k, v_k)$$

$$p(f(\boldsymbol{r})) = \sum_k P(z(\boldsymbol{r}) = k)\, \mathcal{N}(m_k, v_k) \quad \text{Mixture of Gaussians}$$

- Separable iid hidden variables: $p(\boldsymbol{z}) = \prod_{\boldsymbol{r}} p(z(\boldsymbol{r}))$
- Markovian hidden variables: $p(\boldsymbol{z})$ Potts-Markov:

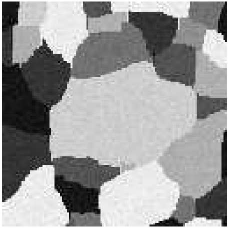
$$p(z(\boldsymbol{r})|z(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})) \propto \exp\left[\gamma \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right]$$

$$p(\boldsymbol{z}) \propto \exp\left[\gamma \sum_{\boldsymbol{r} \in \mathcal{R}} \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right]$$

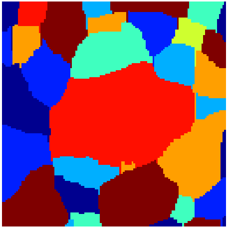
Four different cases

To each pixel of the image is associated 2 variables $f(\boldsymbol{r})$ and $z(\boldsymbol{r})$

- $\boldsymbol{f}|\boldsymbol{z}$ Gaussian iid, $\boldsymbol{z}$ iid :
  Mixture of Gaussians
- $\boldsymbol{f}|\boldsymbol{z}$ Gauss-Markov, $\boldsymbol{z}$ iid :
  Mixture of Gauss-Markov
- $\boldsymbol{f}|\boldsymbol{z}$ Gaussian iid, $\boldsymbol{z}$ Potts-Markov :
  Mixture of Independent Gaussians
  (MIG with Hidden Potts)
- $\boldsymbol{f}|\boldsymbol{z}$ Markov, $\boldsymbol{z}$ Potts-Markov :
  Mixture of Gauss-Markov
  (MGM with hidden Potts)



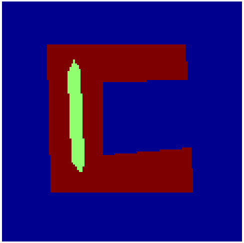
$f(\boldsymbol{r})$

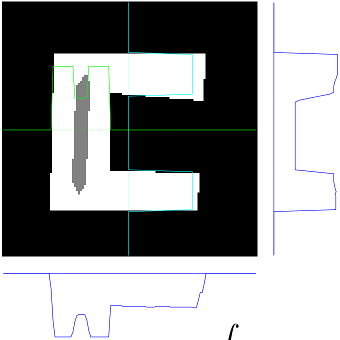


$z(\boldsymbol{r})$

Application of CT in NDT

Reconstruction from only 2 projections



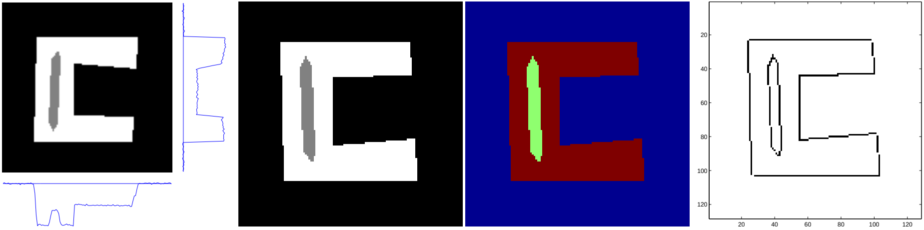


$$g_1(x) = \int f(x,y) \, \mathrm{d}y, \qquad g_2(y) = \int f(x,y) \, \mathrm{d}x$$

- Given the marginals $g_1(x)$ and $g_2(y)$ find the joint distribution $f(x,y)$.
- Infinite number of solutions : $f(x,y) = g_1(x) \, g_2(y) \, \Omega(x,y)$
  $\Omega(x,y)$ is a Copula:

$$\int \Omega(x,y) \, \mathrm{d}x = 1 \quad \text{and} \quad \int \Omega(x,y) \, \mathrm{d}y = 1$$

Application in CT



$\boldsymbol{g}\vert\boldsymbol{f}$	$\boldsymbol{f}\vert\boldsymbol{z}$	$\boldsymbol{z}$	$\boldsymbol{c}$
$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$	iid Gaussian	iid	$c(\boldsymbol{r}) \in \{0,1\}$
$\boldsymbol{g}\vert\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{H}\boldsymbol{f}, \sigma_\epsilon^2 \boldsymbol{I})$	or	or	$1 - \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))$
Gaussian	Gauss-Markov	Potts	binary

Proposed algorithm

$$p(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta} | \boldsymbol{g}) \propto p(\boldsymbol{g} | \boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta})\, p(\boldsymbol{f} | \boldsymbol{z}, \boldsymbol{\theta})\, p(\boldsymbol{\theta})$$

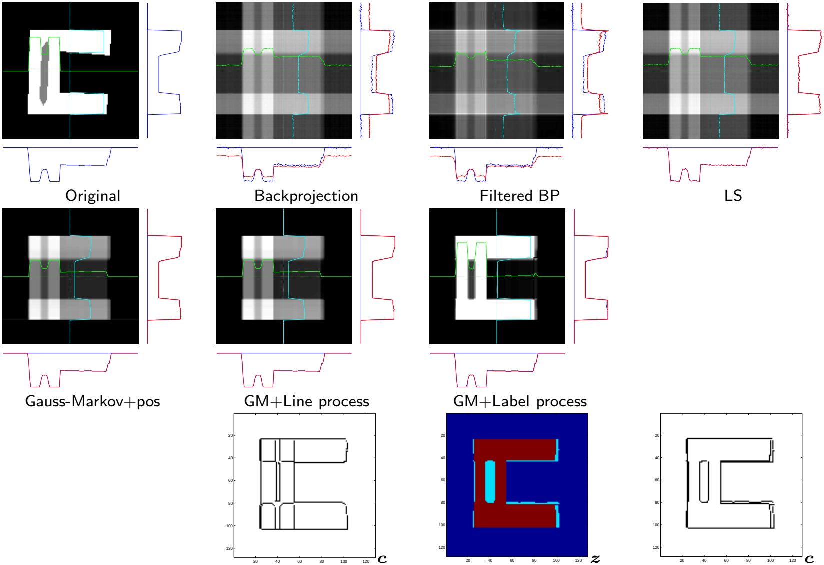
General scheme:

$$\widehat{\boldsymbol{f}} \sim p(\boldsymbol{f} | \widehat{\boldsymbol{z}}, \widehat{\boldsymbol{\theta}}, \boldsymbol{g}) \longrightarrow \widehat{\boldsymbol{z}} \sim p(\boldsymbol{z} | \widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\theta}}, \boldsymbol{g}) \longrightarrow \widehat{\boldsymbol{\theta}} \sim (\boldsymbol{\theta} | \widehat{\boldsymbol{f}}, \widehat{\boldsymbol{z}}, \boldsymbol{g})$$

Iterative algorithme:

- Estimate $\boldsymbol{f}$ using $p(\boldsymbol{f} | \widehat{\boldsymbol{z}}, \widehat{\boldsymbol{\theta}}, \boldsymbol{g}) \propto p(\boldsymbol{g} | \boldsymbol{f}, \boldsymbol{\theta})\, p(\boldsymbol{f} | \widehat{\boldsymbol{z}}, \widehat{\boldsymbol{\theta}})$
  Needs optimisation of a quadratic criterion.
- Estimate $\boldsymbol{z}$ using $p(\boldsymbol{z} | \widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\theta}}, \boldsymbol{g}) \propto p(\boldsymbol{g} | \widehat{\boldsymbol{f}}, \widehat{\boldsymbol{z}}, \widehat{\boldsymbol{\theta}})\, p(\boldsymbol{z})$
  Needs sampling of a Potts Markov field.
- Estimate $\boldsymbol{\theta}$ using
  $p(\boldsymbol{\theta} | \widehat{\boldsymbol{f}}, \widehat{\boldsymbol{z}}, \boldsymbol{g}) \propto p(\boldsymbol{g} | \widehat{\boldsymbol{f}}, \sigma_\epsilon^2 \boldsymbol{I})\, p(\widehat{\boldsymbol{f}} | \widehat{\boldsymbol{z}}, (m_k, v_k))\, p(\boldsymbol{\theta})$
  Conjugate priors $\longrightarrow$ analytical expressions.

Results

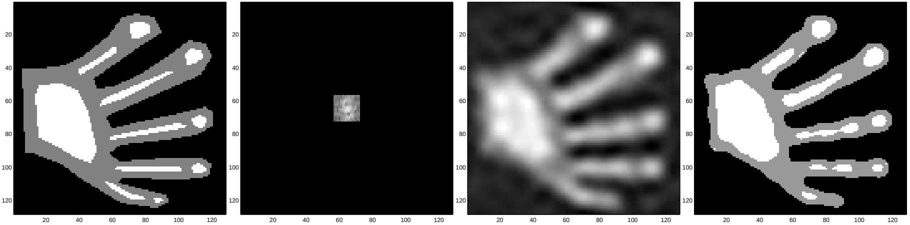

Original
Backprojection
Filtered BP
LS
Gauss-Markov+pos
GM+Line process
GM+Label process
c
z
c

Application in Microwave imaging

$$g(\boldsymbol{\omega}) = \int f(\boldsymbol{r}) \exp\left[-j(\boldsymbol{\omega}.\boldsymbol{r})\right] \, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{\omega})$$

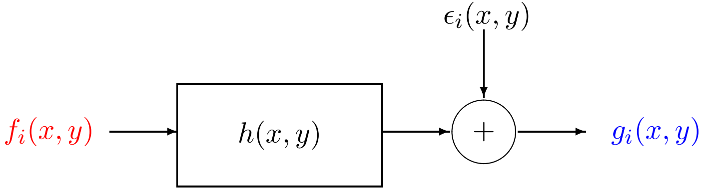
$$g(u,v) = \iint f(x,y) \exp\left[-j(ux+vy)\right] \, \mathrm{d}x \, \mathrm{d}y + \epsilon(u,v)$$

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

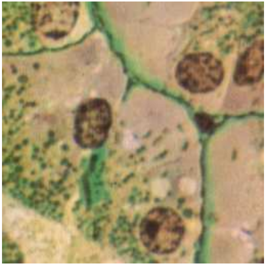


$f(x,y)$ $g(u,v)$ $\widehat{f}$ IFT $\widehat{f}$ Proposed method

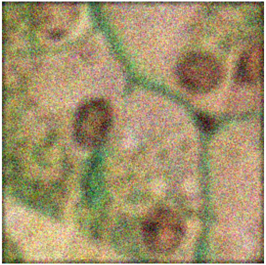
Color (Multi-spectral) image deconvolution



Observation model : $\boldsymbol{g}_i = \boldsymbol{H}\,\boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1,2,3$



$\overset{?}{\Longleftarrow}$

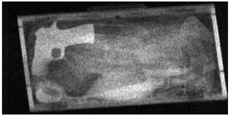


Images fusion and joint segmentation

(with O. Féron)

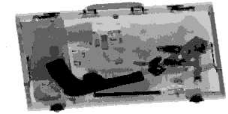
$$
\begin{cases}
g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\
p(f_i(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_{i\,k}, \sigma_{i\,k}^2) \\
p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}) \\
p(\boldsymbol{z}) \propto \exp\left[\gamma \sum_{\boldsymbol{r}\in\mathcal{R}} \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right]
\end{cases}
$$

$\boldsymbol{g}_1$

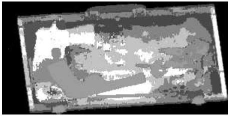


$\boldsymbol{g}_2$

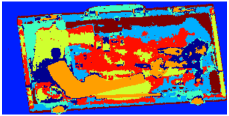
$\longrightarrow$



$\widehat{\boldsymbol{f}}_1$



$\widehat{\boldsymbol{f}}_2$

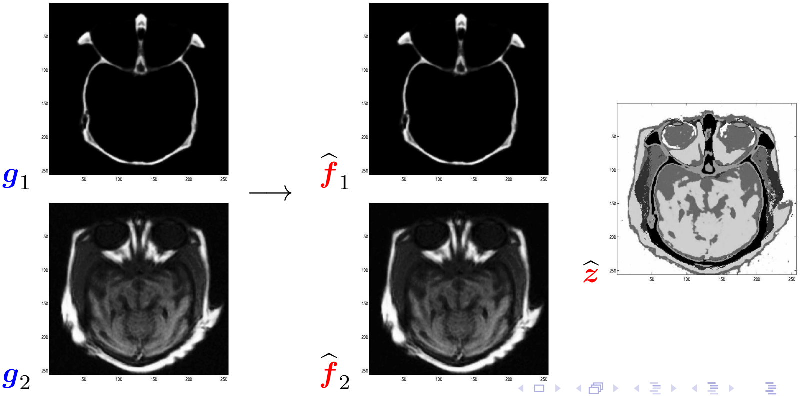


$\widehat{\boldsymbol{z}}$

Data fusion in medical imaging

(with O. Féron)

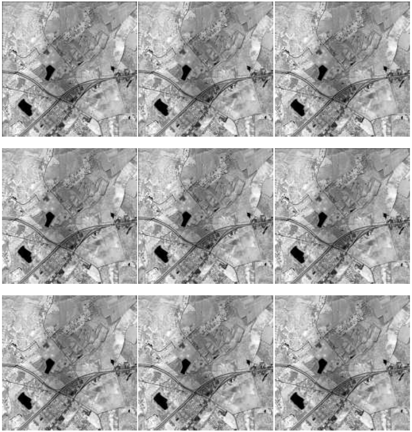
$$\begin{cases} g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_{i\,k}, \sigma^2_{i\,k}) \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}) \\ p(\boldsymbol{z}) \propto \exp\left[\gamma \sum_{\boldsymbol{r}\in\mathcal{R}} \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right] \end{cases}$$

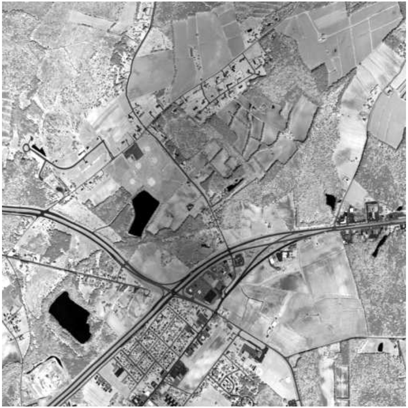


Super-Resolution

(with F. Humblot)

$$\begin{cases} g_i(\boldsymbol{r}) = [\mathcal{D}\mathcal{M}\mathcal{B} f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{i\,k}^2) \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}) \\ p(\boldsymbol{z}) \propto \exp\left[\gamma \sum_{\boldsymbol{r}\in\mathcal{R}} \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right] \end{cases}$$





$\overset{?}{\Longrightarrow}$

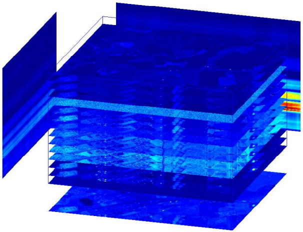
Low Resolution images

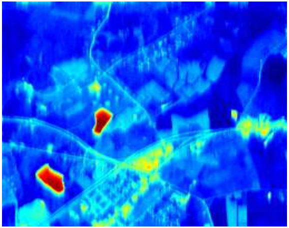
High Resolution image

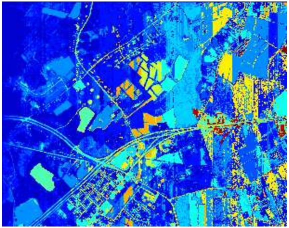
Joint segmentation of hyper-spectral images

(with N. Bali & A. Mohammadpour)

$$\begin{cases} g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_{ik}, \sigma_{i\,k}^2), \quad k = 1, \cdots, K \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}) \\ p(\boldsymbol{z}) \propto \exp\left[\gamma \sum_{\boldsymbol{r}\in\mathcal{R}} \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right] \\ m_{ik} \text{ follow a Markovian model along the index } i \end{cases}$$



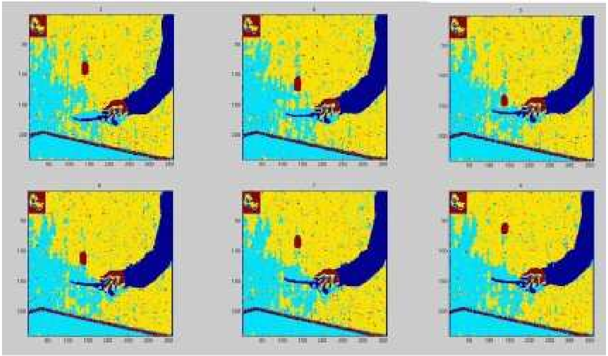




Segmentation of a video sequence of images

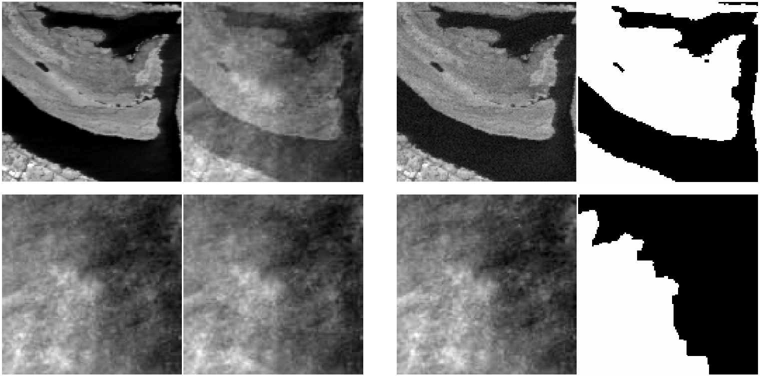
(with P. Brault)

$$\begin{cases} g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z_i(\boldsymbol{r}) = k) = \mathcal{N}(m_{i\,k}, \sigma_{i\,k}^2), \quad k = 1, \cdots, K \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}_i) \\ p(\boldsymbol{z}) \propto \exp\left[\gamma \sum_{\boldsymbol{r}\in\mathcal{R}} \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right] \\ z_i(\boldsymbol{r}) \ \text{follow a Markovian model along the index } i \end{cases}$$



Source separation: (with H. Snoussi & M. Ichir)

$$
\begin{cases}
g_i(\boldsymbol{r}) = \sum_{j=1}^{N} A_{ij} f_j(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\
p(f_j(\boldsymbol{r})|z_j(\boldsymbol{r}) = k) = \mathcal{N}(m_{j\,k}, \sigma_{j\,k}^2) \\
p(\boldsymbol{z}) \propto \exp\left[\gamma \sum_{\boldsymbol{r}\in\mathcal{R}} \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right] \\
p(A_{ij}) = \mathcal{N}(A_{0\,ij}, \sigma_{0\,ij}^2)
\end{cases}
$$



$\boldsymbol{f}$ $\boldsymbol{g}$ $\widehat{\boldsymbol{f}}$ $\widehat{\boldsymbol{z}}$

Conclusions

- Bayesian Inference for inverse problems
- Different prior modeling for signals and images:
  Separable, Markovian, without and with hidden variables
- Sprasity enforcing priors
- Gauss-Markov-Potts models for images incorporating hidden regions and contours
- Two main Bayesian computation tools: MCMC and VBA
- Application in different CT (X ray, Microwaves, PET, SPECT)

Current Projects and Perspectives :

- Efficient implementation in 2D and 3D cases
- Evaluation of performances and comparison between MCMC and VBA methods
- Application to other linear and non linear inverse problems: (PET, SPECT or ultrasound and microwave imaging)

Current Applications and Perspectives

We use these models for inverse problems in different signal and image processing applications such as:

- Period estimation in biological time series
- Signal deconvolution in Proteomic and molecular imaging

- X ray Computed Tomography
- Diffraction Optical Tomography
- Microwave Imaging, Acoustic imaging and sources localization
- Synthetic Aperture Radar (SAR) Imaging

Thanks to:

Graduated PhD students:

1. C. Cai (2013: Multispectral X ray Tomography)
2. N. Chu (2013: Acoustic sources localization)
3. Th. Boulay (2013: Non Cooperative Radar Target Recognition)
4. R. Prenon (2013: Proteomic and Masse Spectrometry)
5. Sh. Zhu (2012: SAR Imaging)
6. D. Fall (2012: Emission Positon Tomography, Non Parametric Bayesian)
7. D. Pougaza (2011: Copula and Tomography)
8. H. Ayasso (2010: Optical Tomography, Variational Bayes)
9. S. Fékih-Salem (2009: 3D X ray Tomography)
10. N. Bali (2007: Hyperspectral imaging)
11. O. Féron (2006: Microwave imaging)
12. F. Humblot (2005: Super-resolution)
13. M. Ichir (2005: Image separation in Wavelet domain)
14. P. Brault (2005: Video segmentation using Wavelet domain)
15. H. Snoussi (2003: Sources separation)
16. Ch. Soussen (2000: Geometrical Tomography)
17. G. Montémont (2000: Detectors, Filtering)
18. H. Carfantan (1998: Microwave imaging)
19. S. Gautier (1996: Gamma ray imaging for NDT)
20. M. Nikolova (1994: Piecewise Gaussian models and GNC)
21. D. Prémel (1992: Eddy current imaging)

Thanks to:

Current PhD students:

- L. Gharsali (Microwave imaging for Cancer detection)
- M. Dumitru (Multivariate time series analysis for biological signals)
- S. AlAli (Electrical imaging of CO2 stocking under the earth)

Master students:

- A. Cai (Non-circular X ray Tomography)
- F. Fuc (Multi component signal analysis for biology applications)

Post-Docs:

- J. Lapuyade (2011: Dimentionality Reduction and multivariate analysis)
- S. Su (2006: Color image separation)
- A. Mohammadpour (2004: HyperSpectral image segmentation)

Thanks my colleagues and collaborators

- B. Duchêne & A. Joisel (L2S) (Inverse scattering and Microwave Imaging)
- N. Gac (L2S) (GPU Implementation)
- Th. Rodet (L2S) (Computed Tomography)
- ————————
- A. Vabre & S. Legoupil (CEA-LIST), (3D X ray Tomography)
- E. Barat (CEA-LIST) (Positon Emission Tomography, Non Parametric Bayesian)
- C. Comtat (SHFJ, CEA) (PET, Spatio-Temporal Brain activity)
- J. Picheral (SSE, Supélec) (Acoustic sources localization)
- D. Blacodon (ONERA) (Acoustic sources separation)
- J. Lagoutte (Thales Air Systems) (Non Cooperative Radar Target Recognition)
- P. Grangeat (LETI, CEA, Grenoble) (Proteomic and Masse Spectrometry)
- F. Lévi (CNRS-INSERM, Hopital Paul Brousse) (Biological rythms and Chronotherapy of Cancer)

References 1

1. A. Mohammad-Djafari, "Bayesian approach with prior models which enforce sparsity in signal and image processing," EURASIP Journal on Advances in Signal Processing, vol. Special issue on Sparse Signal Processing, (2012).
2. A. Mohammad-Djafari (Ed.) Problèmes inverses en imagerie et en vision (Vol. 1 et 2), Hermes-Lavoisier, Traité Signal et Image, IC2, 2009,
3. A. Mohammad-Djafari (Ed.) Inverse Problems in Vision and 3D Tomography, ISTE, Wiley and sons, ISBN: 9781848211728, December 2009, Hardback, 480 pp.
4. A. Mohammad-Djafari, Gauss-Markov-Potts Priors for Images in Computer Tomography Resulting to Joint Optimal Reconstruction and segmentation, International Journal of Tomography & Statistics 11: W09. 76-92, 2008.
5. A Mohammad-Djafari, Super-Resolution : A short review, a new method based on hidden Markov modeling of HR image and future challenges, The Computer Journal doi:10,1093/comjnl/bxn005:, 2008.
6. H. Ayasso and Ali Mohammad-Djafari Joint NDT Image Restoration and Segmentation using Gauss-Markov-Potts Prior Models and Variational Bayesian Computation, IEEE Trans. on Image Processing, TIP-04815-2009.R2, 2010.
7. H. Ayasso, B. Duchene and A. Mohammad-Djafari, Bayesian Inversion for Optical Diffraction Tomography Journal of Modern Optics, 2008.
8. N. Bali and A. Mohammad-Djafari, "Bayesian Approach With Hidden Markov Modeling and Mean Field Approximation for Hyperspectral Data Analysis," IEEE Trans. on Image Processing 17: 2. 217-225 Feb. (2008).
9. H. Snoussi and J. Idier., "Bayesian blind separation of generalized hyperbolic processes in noisy and underdeterminate mixtures," IEEE Trans. on Signal Processing, 2006.

References 2

1. O. Féron, B. Duchène and A. Mohammad-Djafari, Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data, Inverse Problems, 21(6):95-115, Dec 2005.
2. M. Ichir and A. Mohammad-Djafari,
Hidden markov models for blind source separation, IEEE Trans. on Signal Processing, 15(7):1887-1899, Jul 2006.
3. F. Humblot and A. Mohammad-Djafari,
Super-Resolution using Hidden Markov Model and Bayesian Detection Estimation Framework, EURASIP Journal on Applied Signal Processing, Special number on Super-Resolution Imaging: Analysis, Algorithms, and Applications:ID 36971, 16 pages, 2006.
4. O. Féron and A. Mohammad-Djafari,
Image fusion and joint segmentation using an MCMC algorithm, Journal of Electronic Imaging, 14(2):paper no. 023014, Apr 2005.
5. H. Snoussi and A. Mohammad-Djafari,
Fast joint separation and segmentation of mixed images, Journal of Electronic Imaging, 13(2):349-361, April 2004.
6. A. Mohammad-Djafari, J.F. Giovannelli, G. Demoment and J. Idier,
Regularization, maximum entropy and probabilistic methods in mass spectrometry data processing problems, Int. Journal of Mass Spectrometry, 215(1-3):175-193, April 2002.
7. H. Snoussi and A. Mohammad-Djafari, "Estimation of Structured Gaussian Mixtures: The Inverse EM Algorithm," IEEE Trans. on Signal Processing 55: 7. 3185-3191 July (2007).
8. N. Bali and A. Mohammad-Djafari, "A variational Bayesian Algorithm for BSS Problem with Hidden Gauss-Markov Models for the Sources," in: Independent Component Analysis and Signal Separation (ICA 2007) Edited by:M.E. Davies, Ch.J. James, S.A. Abdallah, M.D. Plumbley. 137-144 Springer (LNCS 4666) (2007).

References 3

1. N. Bali and A. Mohammad-Djafari, "Hierarchical Markovian Models for Joint Classification, Segmentation and Data Reduction of Hyperspectral Images" ESANN 2006, September 4-8, Belgium. (2006)
2. M. Ichir and A. Mohammad-Djafari, "Hidden Markov models for wavelet-based blind source separation," IEEE Trans. on Image Processing 15: 7. 1887-1899 July (2005)
3. S. Moussaoui, C. Carteret, D. Brie and A Mohammad-Djafari, "Bayesian analysis of spectral mixture data using Markov Chain Monte Carlo methods sampling," Chemometrics and Intelligent Laboratory Systems 81: 2. 137-148 (2005).
4. H. Snoussi and A. Mohammad-Djafari, "Fast joint separation and segmentation of mixed images" Journal of Electronic Imaging 13: 2. 349-361 April (2004)
5. H. Snoussi and A. Mohammad-Djafari, "Bayesian unsupervised learning for source separation with mixture of Gaussians prior," Journal of VLSI Signal Processing Systems 37: 2/3. 263-279 June/July (2004)
6. F. Su and A. Mohammad-Djafari, "An Hierarchical Markov Random Field Model for Bayesian Blind Image Separation," 27-30 May 2008, Sanya, Hainan, China: International Congress on Image and Signal Processing (CISP 2008).
7. N. Chu, J. Picheral and A. Mohammad-Djafari, "A robust super-resolution approach with sparsity constraint for near-field wideband acoustic imaging," IEEE International Symposium on Signal Processing and Information Technology pp 286–289, Bilbao, Spain, Dec14-17,2011

Current PhD's and projects

PhD's:

1. **Microwave imaging: PhD Leila Gharsalli (co-supervising B. Duchêne)**
2. **Multivariate and multicomponents biological data processing: PhD Mircea Dumitru (co-supervising F. Lévi), ERASYSBIO**
3. **ANR: HONTOMIN, PhD Safa AlAli, (CO2 stock supervising using electrical imaging) (B. Duchne & G. Perruson)**
4. **New methods for reducing dose in Computed Tomography, PhD Li Wang (N. Gac)**
5. **Information fusion for radar target recognition, starting PhD, May Abou Chahine, Thales Systèmes Aéroports**

Post-docs

1. **ANR: SURMITO (Optical imaging), S. Mehrab, (B. Duchêne)**
2. **3D Tomography (SAFRAN), Th. Boulay, (N. Gac)**